Lezioni di Cassandra 626/ Eliza colpisce ancora

(626)—L'alter ego di Cassandra ha fatto questa chiacchierata eretica sulle false IA in un paio di eventi, con una certa soddisfazione…

Lezioni di Cassandra 626/ Eliza colpisce ancora



Figure 1:

(626)—L'alter ego di Cassandra ha fatto una chiacchierata eretica sulle false IA in un paio di eventi pubblici, apparentemente con una certa soddisfazione dei presenti; perché allora non trasformare gli appunti in una esternazione vera e propria, salvandoli dall'oblio digitale?

30 giugno 2025—Se oggi siamo in questa situazione è tutta colpa di Joseph Weizenbaum, noto eretico dell'informatica. Anzi, è colpa di Joseph Weizenbaum e della sua segretaria.

Ma andiamo con ordine.

Nell'ambito dell'intelligenza artificiale, una delle pietre miliari più conosciute è stato ELIZA, un programma informatico, sviluppato appunto da Weizenbaum nel 1966, che ha rivoluzionato la nostra comprensione delle interazioni tra l'uomo e la macchina. Eliza, sebbene rudimentale rispetto agli standard odierni, ha gettato le basi per molte delle tecnologie di chatbot e assistenti virtuali che utilizziamo oggi.

Weizenbaum realizzò la versione originale di Eliza utilizzando SLIP, un linguaggio da lui stesso creato; il programma, lungo poche centinaia di linee, fu poi successivamente riscritto e modificato varie volte in numerosi altri linguaggi.

Per comunicare con Eliza, all'epoca, era necessario utilizzare una tastiera ed un schermo. La segretaria di Weizenbaum fece ovviamente da cavia per la nuova invenzione, e questo fu uno dei momenti più curiosi e significativi nella storia di Eliza. Infatti un giorno, arrivando in ufficio, Weizenbaum scoprì la sua segretaria intenta a conversare, seriamente e di sua iniziativa, con il programma, e fu addirittura invitato a *"lasciarli soli"*, perché stavano parlando di "*questioni intime e personali*".

Questo distrasse Weizenbaum dal suo obbiettivo originale di simulare una conversazione, e lo portò a riflettere invece sulla natura dell'interazione umano-computer; fu colpito infatti dall'"*umanità*" che veniva attribuita al programma; proprio questo lo spinse successivamente ad occuparsi degli aspetti filosofici legati alla creazione dell'Intelligenza artificiale.

Già, l'***Intelligenza Artificiale***, disciplina antica con un nome eccezionalmente efficace e pericoloso. Il termine fu creato nel 1955 da John McCarthy—premio Turing, inventore del LisP, etc. McCarthy morì nel 2011, per sua fortuna prima della "bolla" odierna, che forse avrebbe correttamente definito "follia".

Quando cerca di visualizzare l'IA, a Cassandra viene in mente il volto pensoso di John Wood in "*Wargames*", dove interpreta il professor Steven Falken ("*Salve professor Falken, Strano gioco, l'unica mossa vincente è non giocare. Che ne dice di una bella partita a scacchi?*"). Ricordatevi di questa frase, quando avrete finito questo articolo.

Sono passati 70 anni, 50 da Wargames, e tre anni or sono l'IA ha cominciato a funzionare. Come? **Come una nuova e "migliore" Eliza!**

Non è esatto dire che l'IA abbia cominciato a funzionare solo tre anni fa. In 70 anni di storia sono stati sviluppati i motori di inferenza, le reti neurali, le tecniche di apprendimento profondo, tutte cose che nei loro ambiti funzionano benissimo. Se non li avete mai sentiti nominare, ricorderete certo le notizie che l'IA referta le TAC meglio di un radiologo, e che ha battuto i campioni mondiali di scacchi e Go. La dama europea ancora resiste. Comunque, se vi sembra poco …

Tre anni fa una tecnologia già nota da tempo, gli LLM ("*Large Language Models—Grandi modelli di Linguaggio*"), ha cominciato a funzionare. Perché? Per semplici motivi di scala, cioè l'utilizzo di server più potenti per eseguirli, e di più informazioni con cui allenarli.

La speculazione finanziaria delle dot.com, sempre in cerca di nuove opportunità di far soldi, ci si è buttata a pesce. Improvvisamente chatGPT ed i suoi fratelli sono diventati disponibili a chiunque gratuitamente, ed hanno cominciato ad affascinare tutti, esattamente come Eliza aveva fatto con la segretaria di Weizenbaum.

Ma come Eliza, un LLM non sa niente, non comprende niente, non può rispondere a nessuna domanda, e nemmeno rispondere sempre nello stesso modo.

Gli LLM sanno solo mettere in fila, e molto bene, le parole.

Sono fantastici per manipolare il linguaggio scritto, tradurre da una lingua all'altra, cambiare lo stile, fare riassunti. Ma per il resto, se gli si pongono domande, sanno solo fare come Eliza; sono a tutti gli effetti degli affabulatori di supercazzole.

Gli LLM funzionano su basi simil-statistiche, preparando un database di relazioni fra token (diciamo parole), utilizzando grandi quantità di testo. Successivamente consultano questo database quando gli viene proposta una parola od una frase, restituendo quella che è più probabile come parola successiva, con un calcolo su base "vettoriale".

Giusto per esemplificare, è come se l'LLM avesse trasformato tutte le parole del mondo nell'erba di un prato, con i fili d'erba che puntano tutti in direzioni leggermente diverse. Convertendo anche l'ultima parola della "domanda" in un filo d'erba, la prossima parola è il filo d'erba che nel prato punta nella stessa direzione.

Ripetiamo. Gli LLM sono sistemi a base statistica, non applicano algoritmi. Sono solo bravi a comporre frasi credibili ed appropriate. Sono fatti così. Non possono essere consultati come se fossero "oracoli".

"***Oracolo***" non è una parola qualunque nell'IA, campo dove invece possiede un significato molto preciso. Un **oracolo** è un programma che conosce delle informazioni, e che è in grado di fornire risposte a domande pertinenti il suo settore. Non è come l'Oracolo di Delfi, che a domande rispondeva invece con enigmi!

Ma un LLM certamente non è un oracolo. Gli LLM sono inguaribilmente affetti da mancanza di determinismo; inoltre man mano che l'output che gli viene richiesto si allunga, gli LLM inevitabilmente deviano da quello che ci si aspetterebbe.

Usarli tranquillamente in questo modo suicida è conseguenza del vivere in una "bolla" in cui gli LLM sono ossessivamente descritti e venduti proprio come "oracoli".

Sintetizzando, l'IA fatta di LLM e venduta in questo modo è solo *fuffa*. Perché?

Perché vende un prodotto disonesto fin dall'origine, ed inutili allo scopo per cui viene venduto, cioè dare risposte.

Perché descrive, volutamente ed in maniera fuorviante, il funzionamento dell'IA in termini di *intenzionalità*, con termini sempre *antropomorfizzanti*, invece che descrivendoli solo come potenti strumenti per elaborare il linguaggio.

Antropomorfismo per meglio nascondere che si tratta di generatori di discorsi a caso.

Generatori di discorsi a caso che vengono poi "allineati" utilizzando circuiti censori in ingresso ed in uscita, in modo che non rispondano a domande scomode, e non diano risposte scomode (ho detto volutamente "scomode", non "errate"). Poi anche per tentare di intercettare le "risposte" più pericolose.

Ma gli LLM sono ormai stati *allenati* con tutti i testi esistenti al mondo. In passato il loro output è stato "migliorato" aumentandone le dimensioni.

Ma oggi gli LLM non possono più aumentare in dimensione, semplicemente perché non ci sono altre informazioni da fornirgli. Sostanzialmente, **gli abbiamo già dato da "mangiare" l'intera Infosfera**.

E le nuove informazioni che vengono oggi create non possiamo dargliele, perché sono quasi tutte generate proprio dalle IA. E' infatti dimostrato oltre ogni dubbio che le informazioni scritte dalle IA sono per loro un cibo tossico, che le fa degenerare e collassare.

Ma allora perché siamo in questa situazione? Un complotto plutocratico? Una corsa agli armamenti digitale?

E perché le grandi dot.com (Meta, Google, Microsoft, etc.) stanno infilando chatbot IA dappertutto senza che nessuno glielo richieda, e spesso gratis?

In realtà una spiegazione dello stato attuale dell'IA fatta esclusivamente con gli LLM è possibile esaminandola sul piano finanziario. Parliamo, ad esempio, di OpenAI e Sam Altman.

OpenAI ha fatto un passivo di 5 miliardi nel 2024. Anthropic più o meno lo stesso. Pozzi senza fondo, dove il denaro che potrebbe essere speso per cose ben più utili sparisce senza lasciare traccia. Altman ha recentemente comunicato agli investitori che non farà profitti prima del 2029, e che per allora dovrà spendere altri 40 miliardi. Poi, vedi caso, annuncia la prossima cosa meravigliosa; un ancora misterioso device per l'IA che realizzerà insieme all'ex-designer di Apple Joni Ive.

Queste di OpenAI e quelle equivalenti di Anthropic e degli altri sono azioni di *distrazione di massa*, fatte per continuare a tenere eccitati gli investitori che devono (lo dice la parola stessa) continuare ad investire.

Due ultime cose, le più importanti, per concludere.

La prima: non disperatevi. Potete usare tranquillamente gli LLM per abbellire i vostri testi e tradurre le vostre email (ma rileggete sempre il risultato). In questo sono strumenti potenti ed utilissimi.

Al contrario, non usateli mai per realizzare una bozza di un documento o di una perizia; risparmierete tempo, ma compirete un lento, e talvolta veloce, suicidio mentale e professionale.

La seconda: Cassandra voleva terminare **affrontando la questione di quanto il nostro stesso giudizio sull'utilizzo degli LLM possa essere inconsciamente condizionato proprio dal loro uso.**

Ha però trovato un post di **Baldur Bjarnason**, che su questo argomenta in maniera davvero molto accurata, e certamente migliore di quella di Cassandra, arrivando alle stesse conclusioni. La modestia, oppure la pigrizia, hanno vinto, e perciò ne trovate la traduzione qui sotto.

È quasi impossibile per le persone valutare i benefici o i danni di chatbot e agenti attraverso l'auto-sperimentazione. Questi strumenti innescano una serie di pregiudizi ed effetti che offuscano il nostro giudizio.

I modelli generativi presentano inoltre una volatilità dei risultati e una distribuzione non uniforme dei danni, simili a quelle dei prodotti farmaceutici, il che significa che è impossibile scoprire autonomamente quale sarà il loro impatto sociale o persino organizzativo.

La ricerca in ambito tecnologico, software e produttività è estremamente scarsa e si limita per lo più a mero marketing, spesso replicando le tattiche dei settori dell'omeopatia e della naturopatia. La maggior parte delle persone nel settore tecnologico non ha la formazione necessaria per valutare il rigore o la validità degli studi nel proprio campo. (Potreste non essere d'accordo, ma vi sbagliereste.)

*L'enorme portata della bolla dell'"IA" e la quasi totalità del consenso istituzionale—università, governi, istituzioni—implica che **tutti siano di parte** [anche io e voi, NdT]. Ed anche se non lo siete voi stessi, lo saranno il vostro manager, la vostra organizzazione o i vostri finanziatori.*

Persino coloro che cercano di essere imparziali sono intrappolati in istituzioni che gonfiano la bolla e sentiranno il bisogno di proteggere la propria carriera, anche se solo inconsciamente.

Ancora più importante, non c'è modo per il resto di noi di conoscere l'entità dell'effetto che la bolla ha sui risultati di ogni singolo studio o articolo, quindi dobbiamo presumere che li influenzi tutti. Anche quelli realizzati dai nostri amici. Anche gli amici possono essere di parte.

La bolla significa anche che non ci si può fidare assolutamente della classe dirigente e manageriale. A giudicare dalle precedenti bolle speculative, sia nel settore tecnologico che in quello finanziario, le persone oneste che capiscono cosa sta succedendo ne sono quasi certamente già uscite.

Quando comprendiamo qualcosa solo a metà, chiudiamo il cerchio dall'osservazione alla convinzione affidandoci al giudizio dei nostri pari e delle figure autorevoli, ma questi gruppi nel settore tecnologico sono attualmente quasi certamente in errore o sostanzialmente di parte riguardo ai modelli generativi. Questa è una tecnologia praticamente fatta su misura per essere compresa solo a metà dalla tecnologia in generale. Ne afferrano le basi, forse alcuni dettagli, ma non completamente. La "metà" della loro comprensione lascia uno spazio cognitivo che consente a quella convinzione mal fondata di adattarsi a qualsiasi altra convinzione la persona possa avere e a qualsiasi contesto in cui si trovi senza conflitti.

Combinando questi quattro problemi, abbiamo la ricetta per quella che è di fatto una superstizione omeopatica che si sta diffondendo a macchia d'olio in una comunità in cui tutti si convincono che li renda più sani, più intelligenti, più veloci e più produttivi.

Questo sarebbe un male in qualsiasi circostanza, ma i danni che i modelli generativi arrecano all'istruzione, all'assistenza sanitaria, ai vari servizi sociali, alle industrie creative e persino alla tecnologia (ad esempio, l'eliminazione delle posizioni di programmazione entry-level significa la mancanza di programmatori senior in futuro) si stanno configurando come enormi; i costi per gestire questi specifici tipi di modelli rimangono molto più alti dei ricavi e l'infrastruttura necessaria per realizzarli sta soppiantando i tentativi di transizione energetica in paesi come Irlanda e Islanda.

Se mai è esistita una tecnologia per la quale l'atto razionale e responsabile è stato quello di aspettare e aspettare che la bolla scoppiasse, quella è proprio l'"IA".

Scrivere a Cassandra—Twitter—Mastodon
Videorubrica "Quattro chiacchiere con Cassandra"
Lo Slog (Static Blog) di Cassandra
L'archivio di Cassandra: scuola, formazione e pensiero

Licenza d'utilizzo*: i contenuti di questo articolo, dove non diversamente indicato, sono sotto licenza Creative Commons Attribuzione—Condividi allo stesso modo 4.0 Internazionale (CC BY-SA 4.0),* tutte le informazioni di utilizzo del materiale sono disponibili a questo link.

By Marco A. L. Calamari on June 30, 2025.

Canonical link

Exported from Medium on August 27, 2025.